

Thos. N. Jenkins

SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION

BY

HERBERT WESLEY ROGERS, Ph.D.

ARCHIVES OF PSYCHOLOGY

EDITED BY R. S. WOODWORTH

No. 49

COLUMBIA UNIVERSITY CONTRIBUTIONS TO PHILOSOPHY AND PSYCHOLOGY
VOLUME XXVII, No. 4

NEW YORK

APRIL, 1922

AGENTS: G. E. STECHERT & CO. : London (2 Star Yard, Cary St., W. C.); Paris (16, rue
de Condé)

ARCHIVES OF PSYCHOLOGY

SUB-STATION 84, NEW YORK CITY. The Subscription price is five dollars per volume of 500-600 pages. Volume I comprises Nos. 1-10, Volume II, Nos. 11-18, Volume III, Nos. 19-25, Volume IV, Nos. 26-32, Volume V, Nos. 33-39 and Volume VI, Nos. 40-46. The numbers are as follows:

1. The Psychology of Mentally Deficient children: NAOMI NORSWORTHY. *Out of print.*
2. On the Functions of the Cerebrum: The Frontal Lobes: SHEPHERD IVORY FRANZ. 50c.
3. Empirical Studies in the Theory of Measurements: EDWARD L. THORNDIKE. 50c.
4. Rhythm as a Distinguishing Characteristic of Prose Style: ABRAM LIPSKY. 50c.
5. The Field of Distinct Vision: W. C. RUEDIGER. 70c.
6. The Influence of Bodily Position on Mental Activities: ELMER E. JONES. 50c.
7. A Statistical Study of Literary Merit: FREDERIC LYMAN WELLS. 30c.
8. The Relation between the Magnitude of the Stimulus and the Time of Reaction: SVEN FROEBERG. 35c.
9. The Perceptual Factors in Reading: FRANCIS MARION HAMILTON. 50c.
10. Time in English Verse Rhythm: WARNER BROWN. 70c.
11. The Hearing of Primitive Peoples: FRANK G. BRUNER. \$1.00.
12. Studies in Learning and Learning: EDWIN A. KIRKPATRICK. \$1.00.
13. The Inaccuracy of Movement: H. L. HOLLINGWORTH. 80c.
14. A Quantitative Study of Rhythm: HERBERT WOODROW. 60c.
15. The Psychology of Efficiency: HENRY ALFORD RUGER. *Out of print.*
16. On certain Electrical Processes in the Human Body and their Relation to Emotional Reactions: FREDERIC LYMAN WELLS and ALEXANDER FORBES. 40c.
17. The Relative Merit of Advertisements: EDWARD K. STRONG, JR. \$1.00.
18. Attention and Movement in Reaction Time: J. V. BREITWIESER. 50c. (Cloth, 75c.)
19. An Empirical Study of Certain Tests for Individual Differences: MARY THEODORA WHITLEY. \$1.25. (Cloth, \$1.50.)
20. Visual Acuity with Lights of Different Colors and Intensities: DAVID EDGAR RICE. 50c. (Cloth, 75c.)
21. The Curve of Forgetting: C. H. BEAN. 45c. (Cloth, 70c.)
22. The Influence of Caffein on Mental and Motor Efficiency: H. L. HOLLINGWORTH. \$1.50. (Cloth, \$1.75.)
23. Reaction Time to Retinal Stimulation. A. T. POFFENBERGER, JR. 70c. (Cloth, 95c.)
24. Interference and Adaptability: ARTHUR JEROME CULLER. 75c. (Cloth, \$1.00.)
25. Reaction to Multiple Stimuli: JOHN WELHOFF TODD. 60c. (Cloth, 85c.)
26. A Study in Incidental Memory: GARRY C. MYERS. \$1.00. (Cloth, \$1.25.)
27. A Statistical Study of Eminent Women: CORA SUTTON CASTLE. 80c. (Cloth, \$1.05.)
28. The Mental Capacity of the American Negro: MARION J. MAYO. 60c. (Cloth, 85c.)
29. Experimental Studies in Judgment: H. L. HOLLINGWORTH. \$1.25. (Cloth, \$1.50.)
30. The Psychological Researches of James McKeen Cattell: A Review by Some of His Pupils. \$1.00. (Cloth, \$1.25.)
31. Fatigue and Its Effects upon Control: ISAAC EMERY ASH. 60c. (Cloth, 85c.)
32. The Transfer Effects of Practice in Cancellation Tests: MELVIN ALBERT MARTIN. 60c. (Cloth, 85c.)
33. The Intellectual Status of Children Who are Public Charges: J. L. STENQUIST, E. L. THORNDIKE, and M. R. TRABUE. 50c. (Cloth, 75c.)
34. The Relation of Quickness of Learning to Retentiveness: DARWIN OLIVER LYON. 50c. (Cloth, 75c.)
35. The Overcoming of Distraction and Other Resistances: JOHN J. B. MORGAN. 75c. (Cloth, \$1.00.)
36. The Psychology of the Negro—An Experimental Study: GEORGE OSCAR FERGUSON, JR. \$1.25. (Cloth, \$1.50.)
37. The Effect of Distraction on Reaction Time: JOHN E. EVANS: \$1.00. (Cloth, \$1.25.)
38. The Effect of Humidity on Nervousness and on General Efficiency: LORLE IDA STECHER. 90c. (Cloth, \$1.15.)
39. The Mechanism of Controlled Association: MARK A. MAY. 75c. (Cloth, \$1.00.)
40. Recitation as a Factor in Memorizing: ARTHUR I. GATES. \$1.00. (Cloth, \$1.25.)
41. Mental Fatigue during Continuous Exercise of a Single Function: THOMAS RUSSELL GARTH. 85c. (Cloth, \$1.10.)
42. A Psychological Study of Trade-Mark Infringement: RICHARD H. PAYNTER, JR. 85c. (Cloth, \$1.10.)
43. Individual Differences and Family Resemblances in Animal Behavior: HALSEY J. BAGG. 70c. (Cloth, \$1.00.)
44. Experimental Studies in Recall and Recognition: EDITH MULHALL ACHILLES. 90c. (Cloth, \$1.25.)
45. The Morphologic Aspect of Intelligence: SANTE NACCARATI. 70c. (Cloth, \$1.10.)
46. Psychological Examinations of College Students: F. EDITH CAROTHERS. \$1.25.
47. The Effects of Practice on Judgments of Absolute Pitch: EVELYN GOUGH. \$1.25
48. An Experimental Study of Silent Thinking: RUTH S. CLARK. \$1.40.
49. Some Empirical Tests in Vocational Selection: HERBERT W. ROGERS. 75 cents.
50. Adenoids and Diseased Tonsils: Their Effect on General Intelligence. MARGARET COBB. \$1.00.

This series is a continuation of the psychological part of the *Archives of Philosophy, Psychology, and Scientific Methods*, of which one volume was issued, comprising the following numbers:

- Measurements of Twins: EDWARD L. THORNDIKE. 50 cents.
 Avenarius and the Standpoint of Pure Experience: WENDELL T BUSH. 75 cents.
 The Psychology of Association: FELIX ARNOLD. 50 cents.
 The Psychology of Reading: WALTER FENNO DEARBORN. *Out of print.*
 The Measurement of Variable Quantities: FRANZ BOAS. 50 cents.
 Linguistic Lapses: FREDERIC LYMAN WELLS. \$1.00.
 The Diurnal Course of Efficiency: HOWARD D. MARSH. 90 cents.
 The Time of Perception as a Measure of Differences in Sensations:
 VIVIAN ALLEN CHARLES HENMON. 60 cents.

SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION

BY

HERBERT WESLEY ROGERS, Ph.D.

ARCHIVES OF PSYCHOLOGY

EDITED BY R. S. WOODWORTH

No. 49

COLUMBIA UNIVERSITY CONTRIBUTIONS TO PHILOSOPHY AND PSYCHOLOGY
VOLUME XXVII, No. 4

NEW YORK

APRIL, 1922

AGENTS: G. E. STECHERT & CO. : London (2 Star Yard, Cary St., W. C.); Paris (16, rue
de Condé)

TABLE OF CONTENTS

1. The Problem.	5
2. History.	6
A. Tests for typists.	6
B. Primitive methods in vocational selection.	7
C. Traditional methods in vocational selection.	8
D. Modern methods in vocational selection.	8
a. Vocational miniature.	
b. Vocational sample.	
c. Vocational analogy.	
d. Trade tests.	
e. Miscellaneous empirical tests.	
3. Procedure.	15
A. Materials employed.	15
B. Individuals serving as subjects.	15
a. Group 1.	
b. Groups 2 and 3.	
c. Age, education and experience of groups 1, 2, and 3.	
C. Methods of obtaining measures of abilities.	17
a. Stenography.	
b. Grammar.	
c. Typewriting.	
I. Group 1.	
II. Description of the work of groups 2 and 3.	
III. Group 2.	
IV. Group 3.	
D. Some remarks on the procedure.	21
E. Administering the tests.	22
a. General procedure.	
b. Instructions to the subjects.	
c. Records.	

4. Results and how obtained.	24
A. Scores in the tests.	24
B. Treatment of errors.	25
C. Methods of combining the measures.	25
D. Correlations.	26
a. Stenography and grammar.	
I. Method used.	
II. Correlations.	
b. Typewriting.	
I. Method used.	
II. Correlations.	
E. Discussion of correlations.	27
a. Significant tests for typewriting and significant tests for stenography and grammar.	
b. Typewriting, group 1.	
c. Typewriting, groups 2 and 3.	
d. Typewriting, groups 1, 2, and 3.	
F. Partial correlations.	29
5. Applicability of the results.	30
Appendix.	41
1. The supervisor as a judge of abilities.	41
2. Averages of the scores in the tests arranged according to age and education.	42
3. Sample of an army trade test.	44
4. Bibliography.	45

SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION

1. *The Problem.*

The aim of this investigation is to test out in practice the method of empirical vocational tests and to discover, if possible, significant tests for specific commercial functions.

The type of work selected for correlation with the tests is typewriting. Data in stenography and grammar, obtained with one of the groups of subjects, have been utilized and included in the investigation.

2. *History.*

A. Tests for typists.

Lahy (32)* reports an experiment made with eleven typists as subjects; six women who had had from two to four years practice, and five men who had had from two to eight years practice, in typing. Measures of the following functions or processes were obtained; muscular sensibility of the two hands, speed of auditory reactions, memory for numbers and sentences, attention, and several higher processes such as imagination, abstraction, judgment and reasoning. In comparing the three good and the three mediocre women subjects it was found that the former possessed a finer threshold of tactile and muscular discrimination, a tendency of the hands to become equally strong, a better memory, fewer errors of attention, but showed longer auditory reaction times. In comparing the men subjects the same distinctions were found. This is an interesting and valuable contribution in that it stimulates work along these lines and blazes the trail, but it has not been worked out extensively enough to be practically applicable.

Lough (35) reports graphical curves showing the relation between the time required to perform a letter substitution test, "habit formation test," measured after repeated trials with the test, and ability in stenography and typewriting for a group of high school students. The curves show a good degree of correspondence between performances in the test and judgments of ability in stenography and typewriting. To quote Lough† "It is evident that if we give these pupils a test in habit formation first, we can determine what pupils are capable of doing good work in typewriting, and we can save some of the pupils a great amount of wear and tear, by simply allowing them to take some other work in place of typewriting. They are not adapted to the motor responses which typewriting calls for. By means of this experiment it would be possible to select those who are likely to succeed in typewriting and to give a vocation to those who would not succeed." If this test could accomplish all that Lough claims for it it would indeed be a boon to the human race but unfortunately, the necessary scientific data concerning the experiment are not given. There is only one test used and the

*Refers to Bibliography. †p. 94.

investigator who would build up a system of vocational guidance or selection for such a type of work on the ability to perform any one such test, when the original data are inconclusive, would be incautious indeed. However, a good deal of credit is due Lough for his pioneer effort.

Link (33) gives a series of tests to detect ability in type-writing and a series of tests to detect ability in stenography. The tests are, for typing—letter substitution, Trabue completion, context reading, spelling and typing; for stenography—context reading, mixed letters and numbers, substitution, Trabue completion, spelling, grammar, dictation and typing. Link states* “. . . .relevant tests were given to two senior classes of over three hundred girls and boys in a commercial high school, to twenty-six pupils of two business schools, to a group of twenty-two office typists, to another group of nineteen stenographers, to over four hundred candidates for positions as typists and stenographers. . . .the tests selected on the basis of these experiments are those which showed the highest and most consistent agreement with the abilities of those examined.” These are broad statements, made in the absence of any norms, methods used, correlations, or any reliable data whatsoever, which should be presented if “consistent agreement” is to be proved.

B. Primitive methods in vocational selection.

In primitive magic there is the naive idea that by simply expressing a wish forcibly enough the wish will be materialized; thus if the parents of an Indian boy wished their son to be tall, they would toss him in a blanket. In medieval clairvoyance there was and is still a belief that there are signs which reveal an individual's character. The sources of these signs are principally to be found in the zodiac, palm of the hand, tea leaves, cards, dreams, crystal gazing, etc. In the nineteenth century there was a spot located in the brain for the function of speech and later all parts of the body were found to be localized in the brain. Phrenology was a distortion of this fact. Signs disconnected with the individual's body are now given up and internal and personal external signs take their place. In phrenology there was no comparison of skulls, but individuals were taken at random who had some peculiarity in the formation of their skulls; for instance, if

*p. 422-3 and chapter 8.

an individual had a bump anywhere on his head and was also known to possess some peculiarity such as stinginess, then that bump was labelled stinginess. When it became known that there were no scientific criteria for phrenology, the skull was discarded and the face was taken up, and the pseudo-science of physiognomy was born. Physiognomy maintains that there is something in the features which is significant of inner hidden traits. This is true in so far as the characteristic is an expression of a quality of character. But in vocational psychology we are mainly interested in young people whose characteristics have not been well developed or well stamped upon their faces.

C. Traditional methods in vocational selection.

In the selection of employees the employer, for the most part, has been content to form a ready estimate of a candidate's ability by looking him over, or else has accepted him on the authority of someone who has recommended him; or simpler still he tries the candidate at the task and, sink or swim fashion, he promotes him or discharges him according as he succeeds or fails in the work assigned. None of these methods of measuring a man's value will bear criticism. The last is always costly and time consuming; the second is uncertain, depending on the degree of confidence that one has in the recommender, his ability to evaluate and his motives; as to the first there may be some men who are geniuses at picking men of worth at sight; they seem to have an unerring insight which prompts them to their choice, yet even in such cases, it appears that one does not know the signs of value intuitively but the ability to discriminate has been acquired only through long and costly experience.

D. Modern methods in vocational selection.

a. Vocational miniature.

In the vocational miniature, the work is reproduced on a small scale apparatus which duplicates the actual situation which the worker faces while engaged at his task. Typical of this method is an experiment of McComas of Princeton, in which he attempted to detect good and bad telephone operators. He constructed an actual switchboard on a small scale. The operator, to be tested, made connections at the board which were timed on a kymograph in an adjoining room. This

species of "reaction and co-ordination" time showed the interval between the appearance of a light over a call connection and the moment that the operator "plugged in," and also between the moment that a number was called and the appropriate connection made. Fifty-one trials were made by each of the nine operators tested and the average time of each subject was compared with the average ranking of the subjects by the two supervisors. The test detected the two best and two out of the three poorest operators. The correlation of these rankings with the rankings of the supervisors was .6250 with a probable error of .14. This method has the advantage of concreteness and apparent relevance but there is little actual assurance that the ideas and feelings aroused correspond with the processes aroused in the actual work.

b. Vocational sample.

The method of vocational sampling is closely related to the method of miniatures. It simply consists of taking an actual piece of the work to be performed and sampling the candidate's ability by the degree of his success in the trial. Thus, in testing applicants for jobs as filing clerks, they may be given a number of cards to be arranged in order according to date, number, alphabetical position, address, etc. In certain cases specimens of work have been devised or taken into the psychological laboratory and the worker watched more carefully and measured more exactly. Unfortunately this is a method which cannot be applied to very many kinds of work, for care must be taken to make the tests representative of all the work. At its best this method detects only the presence of ability and is not a means of gauging potential ability.

c. Vocational analogy.

This method consists of an attempt to create, in the laboratory, a situation which arouses and exercises the same mental functions which the actual job arouses and exercises. The material is new, but the attitude and endeavor of the worker are intended to be the same. Munsterberg stated (37) that, by this method, sea captains could be selected who could be relied upon to avoid accidents. The test material consisted of twenty-four cards, the size of playing cards, on which are printed four rows of capital letters namely, A E O U in irregular repetition and interspersed with a few other letters. The person to be tested had to distribute these twenty-four cards as quickly as possible into four piles according

to the numerical preponderance of the four capital letters. The time taken to sort the cards and the number of errors made is taken as the index of ability to safely guide a ship's course. The author of this test admits that there was no correlation scientifically determined between the test and the actual ability to guide a ship's course with safety. The insuperable difficulty with this method is that there is no way of ascertaining that the same mental functions which are called into action on the job are really being exercised and tested in the experiment. This is a highly dangerous method and has done much to put vocational psychology into disrepute in the eyes of men of practical business affairs.

d. trade tests.

The method of trade tests is that of putting specific problems, which are germane to the work to be performed on the job, to the applicant for a job. The answers to the questions give an index of his experience and ability on the job. This idea has been made use of in the past in the examinations of candidates for positions by some concerns and chiefly by some civil service commissions. (Appendix 3 for sample of trade test.)

A good trade test should differentiate between various grades of skill, should produce uniform results in various sections of the country and in the hands of individuals of widely different characteristics and should consume the least amount of time and energy consistent with satisfactory results. While there are all degrees of trade ability among the members of any trade, the method of the trade tests has classified all of the members of any trade into a few groups, usually four, such as novice, apprentice, journeyman, and expert. The novice is a man who has no trade ability whatever, or at least none that could not be paralleled by practically any intelligent man. An apprentice has acquired some of the elements of the trade but is not sufficiently skilled to be entrusted with any important task. The journeyman is qualified to perform almost any work done by members of the trade. An expert can perform quickly and with superior skill any work done by men in the trade. It is essential that the trade test differentiate between the journeyman and the apprentice, and the apprentice and the novice. Trade tests devised to make this classification are of three kinds: oral, picture, and performance. The oral tests are most generally used because they are of low

cost and may be applied to a large number of men in a comparatively short time without much equipment.

The following stages are characteristic of the usual method followed in obtaining a trade test.

1. Investigation and collection of trade data. An investigation is made into the conditions of the trade to determine the following facts. (a) The feasibility of a test in the trade, (in one instance it was found that the trade of gunsmith was not a recognized trade, though there were gun repairers). (b) The elements which require and permit of testing to determine whether men can be graded in the trade according to degrees of skill, (in some trades it has been found that the trade required simply the performance of a single set of operations and there were no gradations among the members of the trade). (c) The kinds of tests that can be used, (some trades such as truck driving are mainly matters of skill and for them performance tests are better than oral or written tests.) Other trades such as interior wiring and power plant operation are mainly matters of knowledge. For these trades oral and picture tests are better. After having discovered by inquiry that the trade is a recognized trade and can be tested, all the information necessary is collected in the field from all available sources, such as experts of the trade, trade union officials, literature of the trade, trade school authorities, employers and the like.

2. Compiling the questions. As a result of the information collected, a number of questions, usually forty to sixty are compiled, each of which calls for an answer which shows knowledge of the trade. The experience of the formulators of trade tests has shown that a good question meets the following requirements. (a) It must be in the language of the trade. (b) It must be a unit, complete in itself and requiring no explanation. (c) It is not a chance question that could be answered by a good guess. (d) It must be as short as possible and must be capable of being answered by a very short answer. (e) It must not be ambiguous.

3. Preliminary sampling. After the large number of questions originally formulated has been sifted down by the application of the requirements listed in the preceding paragraph the questions are used in a preliminary sampling on a number of tradesmen whose answers indicate the merit of the different questions and their grades of difficulty. In this

sampling tradesmen from different shops or plants are tried, in order to guard against specialized methods or modes of expression confined to a single locality.

4. Revision and formulation. The preliminary sampling affords a means of checking on the following points. (a) Is the test applicable to trade conditions? (b) Does the test represent good trade practice? (c) In what way can parts be profitably modified, supplemented or eliminated? (d) Does the test represent the whole range of the trade from the novice to the expert? (e) Is it a representative sampling of the whole range of trade processes? In the light of the answers to these questions the test is revised and then formulated.

5. Final sampling. Final sampling is made by testing a number of men, usually twenty, who are known to be typical members of each group (novice, apprentice, journeyman, expert.) Among the novices tested are usually some highly intelligent and mature men of good general knowledge but no trade ability. Statistical treatment of the results and of the answers to each question enables the determination to be made of a relative value of each individual question and the selection that makes a proper balance.

6. Evaluation. If a trade test is good, a known expert when tested, is able to answer all, or nearly all, the questions correctly; a journeyman is able to answer the majority; an apprentice a smaller part; and a novice practically none. This does not mean that each question should be answered correctly by all the experts, a majority of the journeymen, some apprentices and no novices. There are a few questions which show this result. Other types of questions, however, are more common. Some show a distinct line of cleavage between the novice and the apprentice. Novices fail, but apprentices, journeymen and experts alike answer correctly. There are likewise questions which are answered correctly by nearly all the journeymen and experts, but only a few apprentices and also questions that only an expert can answer. Each type of question has its value in a good test. The main requirement is that the tendency of the curve of distribution of the scores should be upward. A question that is answered correctly by more journeymen than experts or more apprentices than journeymen is undesirable and is at once discarded. A proper balance is made of the others.

7. Calibration. As each question is allowed a number of

points, usually four, it becomes necessary to determine how many points should indicate an expert, how many a journeyman, etc. This is accomplished by noting how many points were scored by the known experts and the known journeymen when they were tested. Ordinarily the expert scores higher than the journeyman and the journeyman higher than the apprentice. It frequently happens that a few journeymen score as high as the lowest of the experts and a few apprentices as high as the lowest of the journeymen. There are, consequently, certain overlappings between the classes. In calibrating, the object is to draw the dividing line between classes so that the overlapping shall be as small as possible. When these dividing lines, or critical scores have been established the test is ready to be applied.

Picture tests are made in practically the same way as the oral tests. The peculiar characteristic of picture tests is that the questions making up the tests relate to illustrations of trade tools and appliances.

Performance tests are fundamentally the same as the method of the vocational sample discussed earlier.

Certain important difficulties of the method of the trade tests may be mentioned. One is due to the fact that trade terms are very fluent, for instance the word "plug" probably means a score of different things in many trades, and probably no single one of these twenty "plugs" is universally known by that name. Localisms exist in various parts of the country such as "come along" in the linesman's trade in New York City. Foreigners, who compose no insignificant part of the total number of tradesmen, have a different nomenclature and one which is not usually a literal translation of the English. There is a difference in trade procedure in different parts of the country, carpenters trained in New York, Georgia and California will answer the question "What is done first in building a house" in very different ways and each in a way which is right according to his training. People may be coached, with ease, upon items of specific information—an individual with the barest trace of literacy can be taught to answer the following written questions: "What is your name?" "In what country were you born?" "In what state in the Union is your domicile?" Coaching is quite as easy in other specific information tests. Good trade tests determine simply the presence of trade experience and the degree of skill actu-

ally possessed. They are not prophetic tests which indicate probable aptitudes for trades which the person tested has never followed.

e. *Miscellaneous empirical tests.*

Finally, there are cases in which tests having vocational significance have been sought by purely haphazard and empirical ways. These may be designated as empirical vocational tests. This method consists of selecting at random a number of tests and obtaining measures of large groups of workers in any vocation and finding out if there are any tests which have a high positive correlation with the actual work. This is a strictly empirical method which is not based upon any apriori postulates as are most of the other methods. In the homely phrase of Carlyle, these tests seem to be "significant of much."

3. Procedure.

A. Materials employed.

Nine tests were used, selected at random, from the group of tests which Woodworth and Wells (53) have standardized for the American Psychological Association. In groups 2 and 3 it would have been unwise to have attempted to use more than this number since it took approximately thirty minutes to test a subject. Under commercial conditions this was a great amount of time for the concern to give, and also for the subjects to give since they were losing their bonus for the time spent on the tests. The tests used are—opposites, verb-object, agent-action, action-agent, color naming, mixed relations, hard directions, number checking and form substitution.* Of those tests in which more than one form is given by Woodworth and Wells the form which was used in this experiment is as follows.

*Considerable time was spent by many subjects on the stimulus words "love" in the opposites test, on "kiss" in the verb-object test, and on "lungs" in the agent-action test. It was evident, that for some subjects, it was more difficult to respond to these words than to others of the list. If the tests were to be given over again these words should be taken out and others, of the same difficulty as the rest of the list, substituted for them.

Opposites	Verb-object	Agent-action	Action-agent	Mixed relations	
north	read	baby	gallops	Eye-see	Ear-
sour	tear	fire	bites	Monday-Tuesday	April-
out	throw	dog	boils	Do-did	See-
weak	paint	laborer	sleeps	Bird-sings	Dogs-
good	mail	pencil	floats	Hour-minute	Minute-
after	light	army	growls	Straw-hat	Leather-
above	sail	heart	sails	Cloud-rain	Sun-
sick	spin	pin	roars	Hammer-tool	Dictionary-
slow	lock	gun	scratches	Uncle-aunt	Brother-
large	wash	eyes	stings	Dog-puppy	Cat-
rich	bake	bird	shoots	Little-less	Much-
dark	spill	wind	melts	Wash-face	Sweep-
front	kiss	lungs	swims	House-room	Book-
love	polish	bell	explodes	Sky-blue	Grass-
tall	sweep	musician	aches	Swim-water	Fly-
open	fill	parrot	glows	Once-one	Twice-
summer	sharpen	clock	news	Cat-fur	Bird-
new	write	ax	cuts	Pan-tin	Table-
come	chew	broom	flies	Buy-sell	Come-
male	drive	mosquito	burns	Oyster-shell	Banana-

B. Individuals serving as subjects.

a. Group 1.

Seventy-seven young men and women, who were studying typewriting, stenography, and grammar in the Extension De-

partment of Columbia University, served as subjects in this first group. These people were divided among four different sections, three evening classes and one day class. In all of the evening classes it was impossible to obtain any adequate measure of the subjects' abilities, for they either dropped out in short time or failed to take the prescribed tests and examinations. The results worked up for the day class of forty-five students, forty-three female and two male are the only ones used in this investigation. The subjects made an intensive study of typewriting and stenography devoting two hours in class to each subject each day for five days a week. Some outside study was also done. The course lasted from the first part of October, 1915, until the middle of May, 1916. In group 1 the performances in the tests were correlated with abilities in typewriting, stenography and grammar.

b. Groups 2 and 3.

One hundred and eighteen typists, all female, in a large retail commercial concern in New York City were tested. Of these ten left the employ of the concern before any adequate measure of their abilities could be determined. The remainder were divided into two groups. Group 2 consisted of thirty-eight typists who had been working in the same division for at least ten months. Group 3 consisted of sixty-five typists who had been working in this same division for at least one month and a half and for not longer than six months. All of the subjects in both groups had been doing the same kind of typing. The records of five typists who had been working between six months and ten months were discarded. This division was made in order to allow, as well as possible, for the effect of practice in the particular kind of typing which the subjects in these two groups were performing, as it would have been unscientific to put all of the subjects in one group.* In groups 2 and 3 the performances in the test were correlated with ability in typewriting only.

c. Age, education and experience of groups 1, 2 and 3.

The figures in the following table are in per cent of the number of individuals in each group.

AGE

	Group 1.	Group 2.	Group 3
16-19 years old.	.46	.68	.91
20-24 "	.32	.32	.08
25-29 "	.06	.00	.01
30-34 "	.11	.00	.00
35-38 "	.04	.00	.00

*See section 3, C, c, 2, (description of the work of groups 2 and 3).

EDUCATION

Grammar school education	.02	.61	.47
1-3 years high school	.02	.36	.47
High School graduates	.51	.03	.05
1-3 years college	.22	.00	.00
College graduates	.22	.00	.00

EXPERIENCE

Never worked previously	.85	.00	.31
Worked at typing 1-5 months	.00	.00	.10
Worked at typing 10-12 months	.00	.66	.26
Worked at typing more than 1 year	.00	.34	.23
Worked at other work than typing	.15	.23	.20

In the age table the mode for all groups is the group 16-19 years old. In the education table the mode for group 1 is the full high school education, for group 2 it is the grammar school education, and for group 3 it is either the grammar school education or the 1-3 years high school class.

C. Methods of obtaining measures of abilities.

a. Stenography.

In February a mid-year examination was given in stenography which was graded by the A, B, C, D, F, method; A, B, C, being the order of satisfactory grades, D a condition and F failed. In each of these grades the mark was further qualified so that the A group, for instance, came to consist of A plus, A, and A minus. The individuals were then ranked in an order of merit in each of the three sub-divisions of each grade, for instance if there were three subjects in the A plus group they were ranked, according to their abilities, 1, 2, and 3. From these last rankings an order of merit series, from one to forty-five, was secured.

b. Grammar.

In February a mid-year examination was given in grammar, spelling, punctuation, paragraphing, letter writing, etc. This was a thorough three hour examination and the results were probably a good index to the subjects' abilities. The

grades in this examination were returned in per cent terms which were readily transformed into an order of merit series.

c. Typewriting.

More objective and reliable grades were obtained in typewriting than in either stenography and grammar, the subjects being measured in actual quantity and quality of output.

I. Group 1.

Each month a typewriting test was given by the instructor, in which the subjects copied a given piece of material on the typewriter, accomplishing as much as they could in ten minutes. The total number of words written in ten minutes, less five words for each error made, divided by ten gave the net number of words written per minute. This is an arbitrary method of scoring adopted as the best relationship between speed and accuracy by the International Typewriting Committee and has been incorporated into the international rules. The following table gives the number of subjects taking the examinations, the average net number of words written per minute, standard deviations, and ranges.

	No. of subjects	Av. no. words per min.	S. D. in words per min.	Range in words per min.
October	42	16.4	6.2	6. -37.
November	40	22.5	7.1	9. -51.
December	40	27.4	6.5	15.6-51.2
February	40	29.7	7.1	15.0-48.1
March	29	34.1	5.0	23.7-44.5
April	27	38.9	4.8	30.4-53.9

II. Description of the work of groups 2 and 3.

The commercial concern, in which the subjects of groups 2 and 3 worked, was composed, in part, of a number of stores which retailed merchandise and which received all of their orders by mail. The customers' orders, in the form received by the concern, were sent to the division in which these subjects worked, to have a sheet typed for each store in the concern from which the customer might have ordered a commodity. Thus an order would require anywhere from one to forty-six sheets written for it, as the customer might order articles, in one letter, from one to forty-six of the stores. A single order contained an order for a commodity from but one store and required but a single sheet to be written for it. Each sheet had typed on it the customer's name and address and specifications of the commodity ordered, such as

amount, size, shape, style, pattern, etc. This was a very particular kind of typing since there were practically no long phrases or sentences written and many figures, symbols and abbreviations were used. Any hierarchy of higher type-writing habits which the subjects might possess appeared to be of little advantage since it could not be made use of directly. All of the subjects in these two groups were working exclusively on the single orders.

It was ascertained, from the total number of words written and the total number of strokes made in several samples of three hundred sheets each that the amount of work done in a day could be expressed in terms of the number of sheets written in the course of a day's work. Thus the number of typed sheets per day furnished the measure of the amount of work turned out by the subjects. A group of "checkers" read each sheet typed by the subjects, compared them with the customer's original order and sent back to the subjects every sheet on which an error was made; and the subjects corrected the error in the course of the day's work, every error being thus penalized. A system was put into effect in which a girl circulated among the subjects and kept each subject supplied with orders, never permitting any subject to run out of orders. By this means no subject was able to select orders and thus receive easy ones.

A weekly bonus on the number of sheets written was in operation in which all of the subjects of group 2 shared almost every week and in which an average of approximately forty per cent of the subjects of group 3 shared. There was much evidence that each girl was urging herself to her best efforts.

III. Group 2.

Records of the average number of sheets written per day were kept for ten months, from these there were selected the records of what appeared to be the best successive five weeks, the selection being based upon; 1. Attendance records; no one subject being absent more than once during this period, in which case she was credited, for the day absent, with the average number of daily sheets that she had written for the rest of the week in which the absence occurred; 2. Season of the year; the season being selected when the business was heaviest and when the subjects were being stimulated, by

their supervisor, to turn out as many sheets as possible. What appeared to be the best week's work was selected in like manner, the best week being a part of the five weeks which were selected. The best day's work was simply that day, within the five weeks' period, in which each individual subject did her best work. The five weeks' measures and the best week's measures were for the same calendar weeks for all subjects, but the best day's measures were not the same for all subjects. Thus three measures of ability were obtained: 1, daily average for five weeks; 2, daily average for one week; 3, best day's performance. The number of subjects, thirty-eight, was constant for each measure. These three cross sections of the subjects' abilities for different lengths of time were selected simply to ascertain whether the tests would correlate to the same degree with each. The following table gives the average number of sheets typed, standard deviations, and range of the average number of sheets typed.

	No. of sheets typed	S. D. of sheets typed	Range of sheets typed
Daily average for five weeks	450.3	64.8	360.2-662.0
Daily average for best week	509.5	68.0	402.5-735.2
Best day's performance	612.0	68.2	531. -812.

IV. Group 3.

Not all of the subjects of this group were working for the concern at the same time. Some subjects entered the employ of the concern two or three months after others and some of these latter had separated from the concern before the former had entered. The constant flux of the coming and going, of the arrival and departure of the individuals in this group, where the labor turnover was over three hundred per cent per year, made it impossible to take any one series of weeks' output, which weeks would be the same for all, as measures of abilities, as was done in group 2. A study of the records showed that many of the subjects began doing very good work about the fifth week after entering the concern. Consequently, the average number of sheets written per day during the fifth and sixth weeks of each individual's term of employment with the concern was taken as the measure of abilities. These weeks, then, were not the same calendar weeks for all individuals. The best day's work was also selected as a measure of abilities, the best day being that

particular day in which each subject did her best work. The following table gives the average number of sheets typed, standard deviations, and ranges. The number of subjects, sixty-five, was constant for both measures.

	No. of sheets typed	S.D.	Range
Daily average for 5th & 6th weeks	311.8	62.3	198.4-420.3
Best day's performance	361.9	65.3	258. -500.

D. Some remarks on the procedure.

The problem of securing groups in which the output of the workers is accurately comparable. In typing it is difficult to obtain a large group of typists in which all of the individuals of the group are performing the same kind of work which can be accurately measured, quantitatively and qualitatively. The difficulty lies in the fact that the group chosen at random is usually a heterogeneous one; that is, one individual does specific work which differs from the specific work of another individual, although both are performing the same generic act—typing. This difficulty has not entered into this investigation since all of the girls were a part of a homogeneous group and all were doing exactly the same kind of typing. The measures obtained in this investigation were not subject to the indeterminable error of a supervisor but were objectively obtained so that it was possible to state that one girl was so many points better or poorer than another girl.

Testing at various periods of the development of practice. In group 1 all of the subjects were tested when they had practically no practice or knowledge of typing. In group 2 all of the subjects were tested after they had acquired considerable practice and skill in typing. In group 3 the subjects were a mixed group being made up of individuals possessing varying degrees of practice and ability in typewriting.

Care taken to secure accurate measures of typing ability. More than one measure of typing ability was used in each group. In group 1 six measures, in group 2 three measures, and in group 3 two measures of ability were used. In group 2 approximately 9100 records of daily output, and in group 3 approximately 3000 records of daily output were tabulated and studied in order to obtain what appeared to be the best measures of typing ability.

Wide mental differences of the groups of subjects. The wide mental differences between group 1 and groups 2 and 3

in educational training indicate that two widely different classes in these respects were tested. It appears that the highest and lowest mental grades of girls in the typing profession were included in this experiment.

E. Administering the tests.

a. General procedure.

The subjects in group 1 were tested within two months after they had started their course. The subjects of groups 2 and 3 were tested after they had acquired considerable speed in typing. All of the tests in groups 2 and 3 and all of the tests but the number checking and the form substitution in group 1 were given individually. The number checking and the form substitution tests in group 1 were given as class tests. The method of conducting the experiment was simple. The subject was seated comfortably at a table, on which the blanks were presented one at a time, after it had been ascertained that the subjects thoroughly understood the instructions. In the number checking, form substitution and hard directions tests replies were filled in in pencil. In the remainder of the tests the replies were spoken. In the tests where the replies were spoken the experimenter, who sat next to the subject, had a copy of each test and if an inaccurate reply was made the subject was required to correct it before giving a reply to the next stimulus word, the subjects being told before hand that they would have to do this. In the number checking test each group was checked which contained any combination of the digits eight and nine.

b. Instructions to the subjects.

In the instructions to the subjects an effort was made to combine the proper comprehension of the experiment by the subject with an ideal uniformity of instruction. In the individual tests instruction was given by description, illustration, and execution. The subject was first clearly told the meaning of the test, then the experimenter performed a small sample of it, and finally the subject herself performed a small sample of the work. These samples were presented in typewritten form and the same samples used with all subjects. Six illustrations of each test were used with all subjects. In the class tests the subjects were simply told what they were to do and the experimenter performed a small sample of the work on the blackboard.

An attempt was made to impress upon the subjects the fact that speed was the main object of the tests. They were given a short talk on the applications of psychological tests and were told that their mental reactions were going to be timed and that their individual results were going to be compared with the results of the rest of the class. They were urged to go as fast as possible but were warned that they would either have to correct their mistakes as they went along or be penalized later for their mistakes (according to whether a class or individual test). It seemed that all the subjects tried to make all the speed possible.

c. Records.

In all the tests in groups 2 and 3 and in all the tests but the two class tests in group 1 the time was taken by a split second watch reading to fifths of a second. The watch was started when the experimenter judged that the subjects' eyes met the first figure or word on the test sheet, which was uncovered by the experimenter. In the class tests instructions were given that as soon as anyone had finished she was to raise her hand. When the first hand was raised the experimenter called upon the class to stop, having instructed them beforehand that he would do this. Thus the time for all was the same and the results were worked up for the amount accomplished.

4. *Results and how obtained.*

A. Scores in the tests.

The following table gives the scores made in the tests for groups 1, 2 and 3.

	Group 1.		Group 2.				Group 3.			
	Av.	S.D.	Av.	S.D.	X	Z	Av.	S.D.	X	Z
Opposites	28.56	8.26	39.00	12.25	.36	.48	43.37	16.12	.52	.95
Verb-object	31.43	7.91	38.59	10.26	.23	.29	37.23	10.39	.18	.31
Agent-action	31.92	10.09	40.61	16.21	.27	.60	39.81	14.28	.25	.41
Action-agent	33.06	6.75	46.33	11.83	.39	.75	50.19	17.35	.51	1.57
Mixed relations	76.03	28.65	99.47	34.30	.31	.20	101.00	34.71	.33	.21
Color naming	53.81	8.99	62.09	8.17	.15	+.09	62.07	9.80	.15	.09
Hard directions	110.01	28.21	186.25	60.00	.69	1.13	188.72	53.91	.71	.91
No. checking	151.52	28.27	157.03	34.72	.04	.23	176.03	33.52	.16	.19
Form subs.	124.40	13.11	150.92	21.70	.21	.65	152.58	30.46	.23	1.32
No. c. Errors	4.70	3.38	8.20	7.62	.74	1.25	7.95	7.83	.68	1.32
Form s. "	.48	.67	2.63	4.48	4.48	3.85	3.17	3.22	5.46	3.81
Hard d. "	1.64	.45	4.33	2.48	1.69	4.51	4.00	2.96	1.44	5.58
No. of subjects	45		38				65			

Av.—average score of the group in seconds and hundredths of a second, in the tests from opposites down through form substitution.* In the last three tests not the time but the average number of errors and hundredths of an error is given.

X—percentage that the average of the group is above or below the average of group 1. A + quantity indicates that the average is below the average of group 1, or a better performance, and no sign before the figures indicates that the average is above the average of group 1, or a poorer performance.

Z—percentage that the standard deviation of the group is above or below the standard deviation of group 1.

*Since in group 1 the grades in the number checking and form substitution tests are measures in terms of the amount accomplished (average in number checking—117.31, average in form substitution—76.31) the figures, in the table under group 1, for these two tests, are roughly derived, for means of comparison in this table only, by dividing the time (constant for the group in each case) by the amount done and multiplying by the number of groups of figures in the number checking test, and the number of figures in the form substitution test.

$$\frac{\text{time}}{\text{amt. done}} \text{ no. of cases.}$$

Groups 2 and 3 showed themselves to be greatly inferior to group 1 in performing the tests. Groups 2 and 3 were made up of girls who were working long hours for very low wages

and who came from families in poor circumstances. Twenty-four per cent of the subjects of group 2 and twelve per cent of the subjects of group 3 came from homes where a foreign language was spoken. The average foreign-born performance, in each test, in each group, was compared with the average performance of the rest of its group and was found to be higher by no more than six per cent in any one test. The inferiority in performing the tests in comparison with group 1 could, therefore, not be traced to difficulty in using the English language due to speaking a foreign language in the home.

B. Treatment of errors.

At the outset of the experiment it was planned to treat separately the time required to perform the test and the number of errors made and to correlate the number of errors made in the number checking, form substitution, and hard directions tests with the grades in stenography, grammar, and typewriting. The standard deviations of the number of errors made in each of the tests being in every case almost as high if not higher, than the average number of errors made, no attempt to correlate the errors with the work has been made since the number of errors made is so unreliable as a measure.

C. Methods of combining the measures.

When several tests have been made of an individual's abilities it is often desirable to show the success of the individual in the series of tests taken as a whole. A good way of doing this is by the order of merit method. The measures of the individuals are so arranged that it can be stated that an individual stood thirteenth from the top in one test, twenty-third in another test, eighteenth in still another test and that her average rank was eighteenth. This method was used in correlating performances in the tests with ability in stenography and grammar.

Although the order of merit method is rather a rough method it has worked well with certain kinds of material in the past. But to transmute a series of quantitative measures into an order of merit series is to throw away a great deal of information contained in the series. Another method consists in taking the average of the group as zero and expressing the individual's standing as a devia-

tion above or below this average. The measure of deviation, usually the standard deviation, is taken as the unit and all deviations are expressed as multiples of this unit. What this method does is to assign to each individual a quantitative position in the distribution of the group. This method was used in getting reduced measures in the tests which were correlated with ability in typewriting in groups 2 and 3 and with grades in typewriting in group 1.

D. Correlations.

a. Stenography and grammar.

I. Method used.

The rank differences formula was used in correlating the tests with abilities in stenography and grammar,

$$r=1-\frac{6\sum d^2}{n(n^2-1)}$$

II. Correlations.

The following table gives the correlations with the tests and abilities in stenography and grammar.

	Steno.	Grammar
Verb-object36	.37
Number checking07	.22
Color naming34	.38
Action-agent23	.35
Agent-action19	.37
Form substitution40	.16
Hard directions46	.54
Mixed relations31	.43
Opposites45	.40
No. of subjects	45	45

b. Typewriting.

I. Method used.

In correlating the tests with abilities in typewriting, in the three groups, a derivation of the Pearson formula was used (52), $r=2Av\left(\frac{a+b}{2}\right)^2-1$ a and b being the reduced measures in the test and typewriting. The following table gives the correlations of the tests with abilities in typewriting.

II. Correlations.

	Group 1.							Group 2.			Group 3.	
	October	November	December	February	March	April	Average Oct. Nov., Dec., Feb.	Five weeks	Best week	Best day	5th & 6th wks.	Best day
Verb-object	.41	.43	.46	.55	.21	.57	.46	.28	.32	.34	.33	.35
No. check.	.45	.47	.37	.53	-.01	.30	.45	.28	.34	.41	.34	.39
Color naming	.30	.43	.55	.41	.29	.61	.40	.39	.39	.40	.36	.38
Action-agent	.42	.43	.29	.41	.00	.40	.39	.13	.29	.26	.29	.28
Agent-action	.29	.19	.40	.31	-.02	.28	.30	.02	.20	.24	.18	.26
Form subs.	.21	.27	.11	.37	.30	.42	.24	.04	.14	.17	.17	.20
Hard dir.	.11	.14	.19	.13	.34	.32	.14	-.20	-.20	-.08	.16	.19
Mixed rel.	-.09	.21	.17	.25	.04	.00	.13	-.07	-.12	-.09	.03	-.04
Opposites	.17	.11	.07	.15	.19	.54	.12	-.09	-.09	-.11	-.07	.06
No. of subj.	42	40	40	40	29	27	40	38	38	38	65	65

E. Discussion of correlations.

a. Significant tests for typewriting and significant tests for stenography and grammar.

In comparing the correlations obtained between tests and typewriting with the correlations between tests and stenography and grammar, if a line is drawn horizontally across the paper in the middle of the list of tests, in each of the preceding tables, it will be observed that, in most cases, the tests which correlated best with stenography and grammar and worst with typewriting lie below this line and the tests which correlate best with typewriting and worst with stenography and grammar lie above this line. The two tests which are not language tests, number checking and form substitution, do not correlate well, except in one instance, with stenography and grammar, but one correlates very well and the other fairly well with typewriting.

b. Typewriting, group 1.

In considering the first four months' correlations in typewriting one is met with the uniformity of the results for any one test during those four months. If a test correlates well with one month's measures of abilities in typewriting, it can be predicted that it will correlate well with another month's measures; if a test correlates poorly with one month's measures it can be predicted that it will correlate poorly with another month's measures.

The March correlations with the verb-object, number checking, agent-action and hard directions tests show no correspondence with the correlations for these tests for the first four months. This can be accounted for as follows: The March typewriting test was an especially difficult one which probably means that it called for reactions which the subjects had not formed the habits of making or that it called for the functioning of processes which did not exist or which were not well developed. Since the exercise was exceptional it could be expected that the results of the exercise would be exceptional and it is quite possible that the correlations would be exceptional and that the uniformity of correlation would be absent. It will also be observed that the number of subjects dropped from forty to twenty-nine this month.

The April correlations show a good degree of correspondence with the correlations for the first four months except for the form substitution, hard directions, and opposites tests. Here again it will be observed that only twenty-seven subjects took the exercise. This reduction in the number of subjects taking the exercise does not mean entirely a process of the elimination of the poorer typists; some of the the poorer typists had dropped out of the course and some of the best typists were in the University offices working as typists. Inspection of the table will show that in the first three tests in the list, those which correlate highest with typewriting, and correlations are usually higher each month as the subjects acquire more speed in typing.

c. Typewriting, groups 2 and 3.

In the tests which correlate well in group 2 the correlations with the best week's measure of ability are higher than the correlations with the five weeks' measure of ability, and the correlations with the best day's measure of ability are higher than either the correlations of the five weeks' or the best week's measure of ability. In the same tests in group 3 the correlations with the best day's measure of ability are higher than correlations with the fifth and sixth weeks' measure of ability. The best day measures the ability when a spurt is made and the best spurt over a long period of time is probably the best index of a subject's ultimate ability that can be found in practice of groups of this sort. Since the correlations are higher in these short intense periods it is concluded

that the tests tend to measure the physiological ability in typewriting.

d. Typewriting, groups 1, 2 and 3.

The seventh column in the table, the average of the correlations for October, November, December and February was obtained as a means of comparison of group 1 with groups 2 and 3 rather than comparing any one month's correlations of group 1 with the correlations of group 2 and 3. In groups 2 and 3 the correlations, for the first four tests in the list, in every case, are lower than those of group 1. With slight exceptions the tests which are correlated well in one group are correlated well in other groups and the tests which are correlated poorly in one group are correlated poorly in other groups.

F. Partial correlations.

The four tests which were correlated highest with typewriting ability were used in deriving a multiple correlation with the measures obtained in the typewriting exercise of February in group 1.

The calculation of the regression equation with five variables is,—Measure in typewriting = .3721 (measure in verb-object) + .0109 (measure in color naming) + .0848 (measure in number checking) + .0852 (measure in action-agent). Or for practical purposes, after multiplying by 24.3, a convenient constant; Measure in typewriting = 9 (measure in verb-object) + .3 (measure in color naming) + 2. (measure in number checking) + 2 (measure in action-agent). After performing the arithmetic this gives a correlation of .61 with a probable error of .07.

5. *Applicability of the results.*

The correlation method has been used to measure the degree of association between performance in the tests and typing ability and it is now necessary to obtain the line of regression or the line of best fit in order to predict, for any individual in the group, her most probable performance in typing from her performance in the tests.

The accompanying diagrams, appended hereto, show frequency distributions in groups 1, 2 and 3 of the reduced measures in the combination of the four tests, as obtained by the weights assigned by the method of partial correlation, and the reduced measures in typewriting, the former being plotted on the x axis and the latter on the y axis. From the general sweep of the scatter diagrams, it is evident that, as the measures in typewriting increase or decrease the measures in the combination of the four tests increase or decrease. The straight line AB, giving the best fit to the data, is fitted to the points in the diagrams by means of the formula $y=mx+b$ which for group 1 becomes $y = .68x + .00b$.*

The expression best fit is seldom defined. Its significance varies with the problem in hand and it generally means a fit which is convenient and which for the problem to be solved, gives satisfactory results. The principle upon which the values of m and b are determined is so to choose m and b as to make the mean square deviation of the observations from the resulting straight line a minimum. The pertinancy of this principle for the problem of prediction is plain since when observations are distributed according to the normal law, the Tables of the Probability Integral enable us to com-

* $y = mx + b$ or $y - \bar{y} = r \frac{\sigma_y}{\sigma_x} x - \bar{x}$ where \bar{y} , \bar{x} and σ_y , σ_x are the means and the standard deviations, respectively of the y's and x's.

$$y - .00 = .61 \frac{109.5}{96} x - .00b.$$

$$y = .68x + .00b.$$

pute the probability of a deviation equal to any multiple or submultiple of the root mean square deviation. Moreover, as in all problems of prediction it is desirable to have the root mean square deviation as small as possible it is obvious that a straight line which fits given data so as to make the mean square deviation of the points from a straight line a minimum is, for the problem of predicting one variable from a knowledge of the other, a good fit to the data.

We have found the best fitting straight line connecting performances in the tests and ability in typewriting by means of the formula $y=mx+b$. For any given value of x , representing a performance in the tests, we can predict, by means of this formula, the most probable performance which will be achieved in typing. The problem is not completely solved by finding the equation connecting the two variables. We are not content to predict the most probable values of y , but we wish to know, in addition, the degree of accuracy of the predictions. The scatter formula supplies an answer to this latter question. The formula S (scatter) $= \sigma_y \sqrt{1-r^2}$ (30)* shows the accuracy with which, using the equation to the best fitting straight line, the magnitude of the one variable may be predicted from a knowledge of the other. Substituting the values, for group 1, in the formula gives $S=84$.† When the association between the two variables is perfect, or when $r = \pm 1$, $S = 0$ and all of the points lie on the straight line, and by means of the equation to the best fitting straight line, we can predict exactly the value of y corresponding to a given value of x . But it is an exceedingly rare case when $r = \pm 1$, and when r lies between these two limiting values, we can still predict the results with a knowledge of the probabilities in favor of the prediction. When the association between the two variables is not perfect, r falls between the limiting values ± 1 , and the equation to the best fitting straight line enables the computation of the most probable value of y corresponding to a given value of x and the value S shows the accuracy with which, using this equation to the best fitting straight line, the magnitude of the one variable may be predicted from a knowledge of the other or enables us to say within what limits any proportion of the measures are scattered about the straight line. The degree of association between performance in the tests and typing ability is measured by the correlation formula whose value is $r = .61$: the formula by means of which typing ability can be predicted from performance in the tests is $y=.68x+.00b$: and the error of the predictions is measured by means of the scatter formula the value of which in this case is $S = 84$.

It is known from the Table of the Probability Integral that when the distribution of the points about the straight line

*For the derivation of this formula the reader is referred to Chapter 2.

† $S = \sigma_y \sqrt{1-r^2} = 106.5 \sqrt{1-.61^2} = 84$.

is normal, 99.7 per cent of the observations lie within a deviation from the straight line equal to $\pm 3S$, 95 per cent lie between $\pm 2S$, and 68 per cent lie between $\pm S$. Since $S = 84$, $2S = 168$. In the chart for group 1 the lines $\pm 2S$ are drawn parallel to the line AB and 168 points above and below it, respectively, on the y axis.

The accuracy of the forecasts is measured by $S = \sigma_y \sqrt{1-r^2}$ where σ_y is the standard deviation of a concrete series, and r is the correlation between the two series. S measures the accuracy of the predictions because it shows how the prediction formula enables us to reduce their variability. If there were no predicting formula the variability of the series that we wish to know would be σ_y but by the use of the formula the variability of the predictions is only $\sigma_y \sqrt{1-r^2}$. The factor $\sqrt{1-r^2}$ measures the reduction in variability that is gained by means of the predicting formula. If, therefore, we wish to compare the accuracy of predictions of two different series, σ_y being constant, the measure of the relative accuracy is given by $\sqrt{1-r^2}$ and the smaller the value of $\sqrt{1-r^2}$, the greater the accuracy of the predictions. The same idea may be put in a different way by saying that the greater the value of r , the greater the accuracy of the predictions.

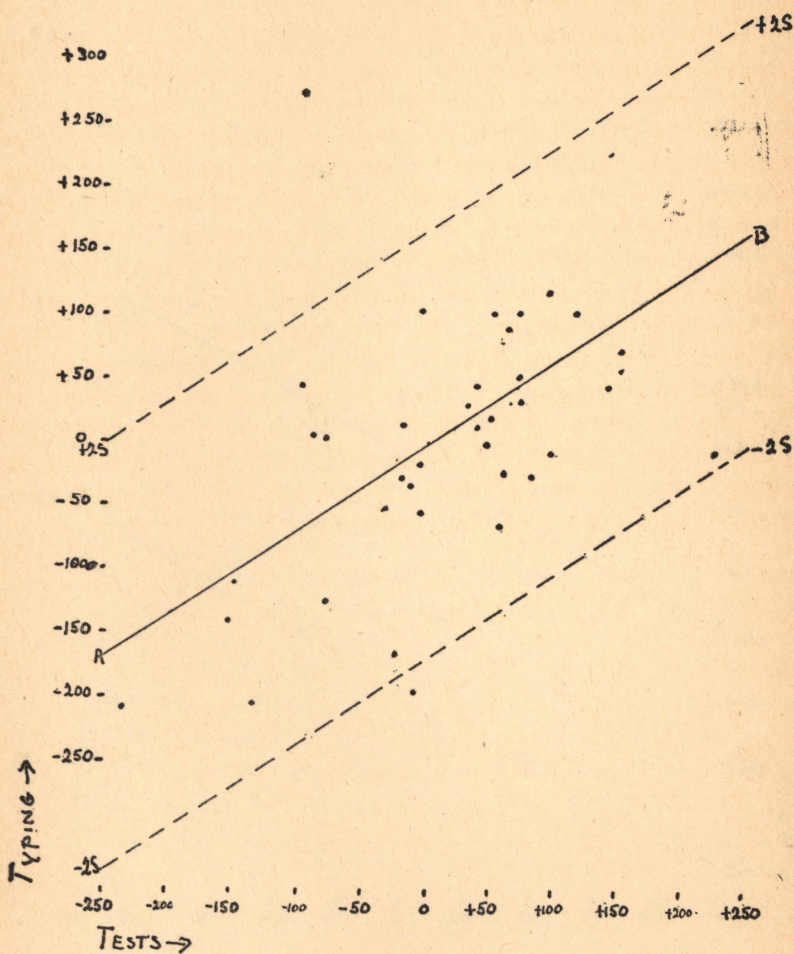


Fig. 1

Scatter Diagram for Group 1. $r = .61$. $y = .676x + .00b$.
 $2s = 168$. Pop. = 40.

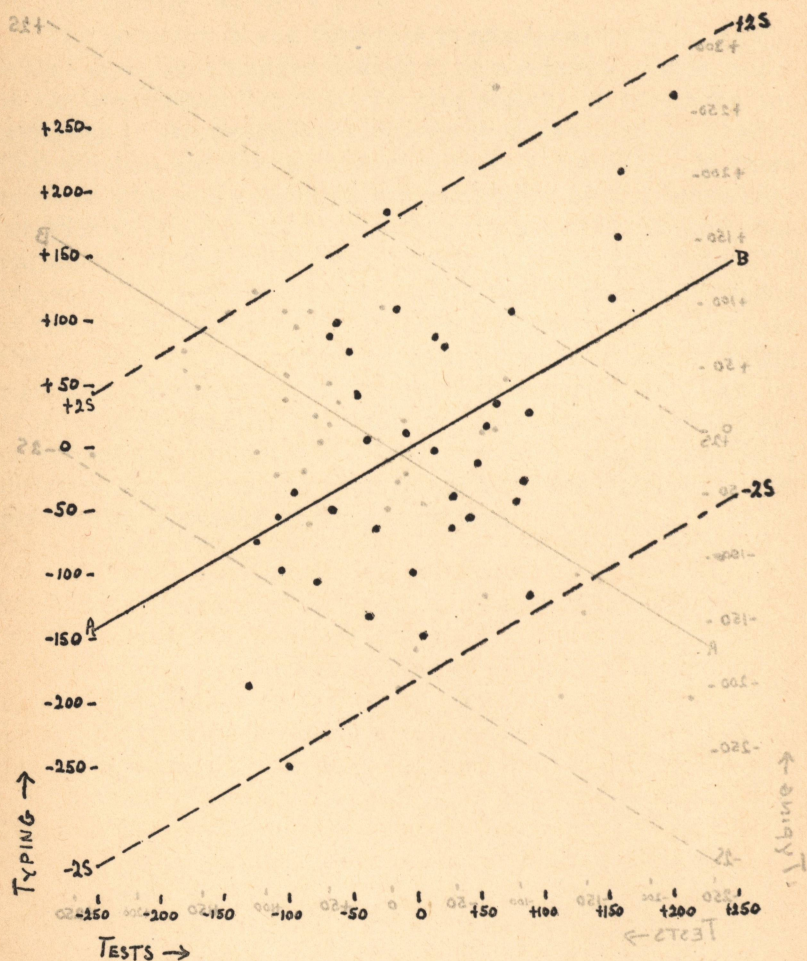


Fig. 21

Scatter Diagram for Group 2. $r = .47$. $y = .59x + .00b$.
 $2s = 186$. Pop. = 38.

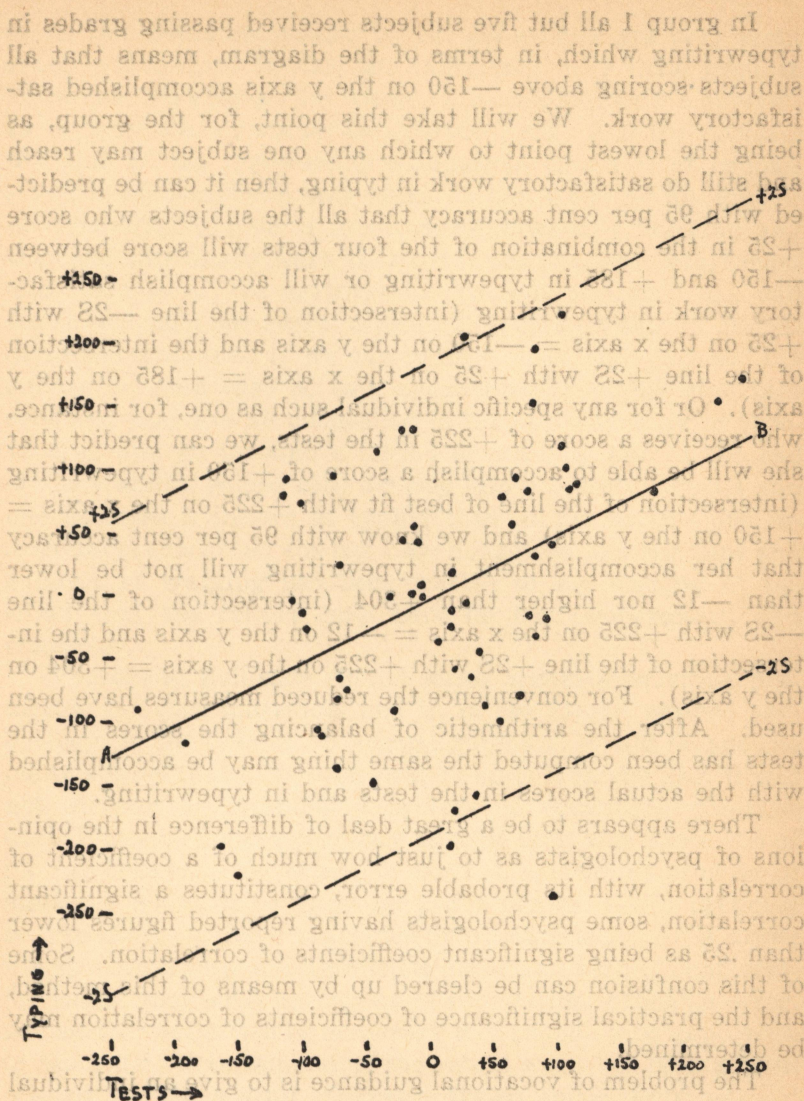


Fig. 3

Scatter Diagram for Group 3. $r = .45$. $y = .515x + .00b$.
 $2s = 187$. Pop. = 65.

In group 1 all but five subjects received passing grades in typewriting which, in terms of the diagram, means that all subjects scoring above -150 on the y axis accomplished satisfactory work. We will take this point, for the group, as being the lowest point to which any one subject may reach and still do satisfactory work in typing, then it can be predicted with 95 per cent accuracy that all the subjects who score $+25$ in the combination of the four tests will score between -150 and $+185$ in typewriting or will accomplish satisfactory work in typewriting (intersection of the line $-2S$ with $+25$ on the x axis $= -150$ on the y axis and the intersection of the line $+2S$ with $+25$ on the x axis $= +185$ on the y axis). Or for any specific individual such as one, for instance, who receives a score of $+225$ in the tests, we can predict that she will be able to accomplish a score of $+150$ in typewriting (intersection of the line of best fit with $+225$ on the x axis $= +150$ on the y axis) and we know with 95 per cent accuracy that her accomplishment in typewriting will not be lower than -12 nor higher than $+304$ (intersection of the line $-2S$ with $+225$ on the x axis $= -12$ on the y axis and the intersection of the line $+2S$ with $+225$ on the y axis $= +304$ on the y axis). For convenience the reduced measures have been used. After the arithmetic of balancing the scores in the tests has been computed the same thing may be accomplished with the actual scores in the tests and in typewriting.

There appears to be a great deal of difference in the opinions of psychologists as to just how much of a coefficient of correlation, with its probable error, constitutes a significant correlation, some psychologists having reported figures lower than .25 as being significant coefficients of correlation. Some of this confusion can be cleared up by means of this method, and the practical significance of coefficients of correlation may be determined.

The problem of vocational guidance is to give an individual counsel as to what sort of an occupation or profession his potential capacities fit him for, so that he can be definitely informed that he will be successful in certain lines of endeavor and unsuccessful in certain other lines, or in more specific cases for any one occupation such as assembling, typewriting, teaching psychology, etc., whether the individual will be a success or a failure and to just what degree he will be a success or a failure in that specific occupation. Unless the chances

of an individual succeeding or not succeeding are very accurately determined it does the individual very little good to know that his chances for success are such and such a percent, as seventy-nine or eighty-two percent, etc. For the problem in hand, in the typewriting occupation, this sort of counsel given to an individual may result in her spending an amount of time ranging from a few months upward in training and then she may finally learn that she is one of those unfortunate ones outside of the seventy-nine percent and that she will never become a successful typist. It is evident that of ten individuals examined if we can predict that the chances for success in each case is, say, approximately eighty-nine percent and they are all counselled to take up the study of typewriting and go through a period of training and at the end of the training period nine develop into successful typists and one becomes a failure at typing then as far as this one individual is concerned she has been misguided and her time and expense of training has been wasted to say nothing of the effect that the failure may have on her temperament and character.

Vocational guidance should mean accurate scientific prediction for each individual and if, in the future, applied psychology, in the field of vocational guidance, would escape the many abuses that have been heaped upon it in the past it will not attempt to establish a system of vocational guidance until it can be almost perfectly sure of each and every one of its predictions. To establish such a system of vocational guidance which is to make no grave errors it will probably be necessary to have a coefficient of correlation well above ninety with a very small probable error and with a small scatter value; just what coefficient of correlation, probable error, and scatter value will be necessary will probably be impossible to accurately determine until the actual work has been accomplished and the scatter diagrams drawn and interpreted.

From the results of this research, it is interesting to know that we can predict that individual A who scores 140 in the psychological tests will most probably attain an ultimate ability in typing of forty words per minute, and in 99.7 percent of the cases she will attain an ultimate ability in typing varying between not more than thirty-one to forty-nine words per minute, and that individual B, who scores ninety-six in the psychological tests will probably attain an ultimate ability in

typing of thirty-two words per minute and in 99.7 percent of all such cases she will attain an ultimate ability in typing varying not more than between twenty-three to forty-one words per minute. But it will readily be seen that these predictions are of very little practical value for the individual in that they afford her very little definite information about just what ultimate ability she, individually, will attain. From our data we can readily predict that in all probability A will become a better typist than B but still it is quite possible for A to develop into a poorer typist than B. The results of this research show, then, that the correlations are not high enough and the scatter value is not small enough to warrant the establishment of a system of vocational guidance whereby an individual can be advised, with any sufficient degree of safety, to undertake or not to undertake to prepare herself to become a typist.

The problem of vocational guidance, then, is supplied by the individual who wants to know of what his potential capacities consist and what sorts of occupations he is fitted to follow successfully and what sorts of occupations he is not fitted to follow successfully. The problem of vocational selection is of an other sort, here the need is supplied by the profession or the industrial concern and not directly by the individual. The problem of vocational selection is—given a certain number of applicants in excess of the number required for a specific job, to select those who, *on the whole*, are best fitted to succeed. In practical life the employment manager has a number of vacancies open for a specific job and of a limited number of applicants for that job his problem is to select from the applicants those who, on the whole, are most likely to succeed. If ten jobs in the typewriting department are open and twenty typists are induced to apply for the jobs the problem is to select the ten who, on the whole, will do the best work. If this can be accomplished there will be an ideal state of affairs. However, this ideal has never been reached in any system of vocational selection. If we have something which works with some degree of success, with such a degree of success never before attained by any other method, which has been checked up scientifically, then we are justified from a pragmatic view point in using it until a better method appears. If we can have a system of vocational selection which can pick from the twenty applicants (ten poor ones and ten good ones)

a majority of the good ones and weed out a majority of the poor ones we have something which works and which is valuable. If by our method of selection of ten typists from twenty applicants eight finally turn out to be very good typists, one a typist of medium ability, and one a poor typist then we have erred to some extent but on the whole we have made a selection which is a good one. By pure chance we would probably have selected five good typists and five poor typists and practically all of the traditional methods of selecting typists appear to do nothing more than this even if they accomplish this much.

It is evident from the foregoing data that this method of empirical vocational tests gives a far more reliable criterion of vocational selection in the field of typewriting than has been obtained by any other method. If a commercial concern finds it necessary to employ a number of typists a test in typing may be given, as so many concerns do, which may or may not determine the actual ability in typing at the moment. This will insure the concern that they are not hiring people who, at the moment are wholly inefficient. However, if we are looking forward to the ultimate ability which the typists will achieve it is not safe to hire a person to do typing on the basis of the performance of a sample of typing for the typing ability of a person when she first enters the employ of a concern is often no index of the ability that she may develop. Along with the test in typing the team of tests may be administered and from the results the employer may predict with 95 percent accuracy or with 99.7 percent accuracy, or with any other degree of accuracy that is required whether or not the applicants will develop into good typists and in addition within just what limits their capacities in typing will enable them to develop in this field, these predictions being made in definite quantitative terms such as the net number of words written per minute.* The prediction of ability by means of the tests will do injustice to some individuals but it will vindicate it-

†In group 3 there was a coefficient of correlation of .28 with a probable error of .10 between the first day's output and the best day's output, the latter measure being taken within a period of six weeks.

*The supply of typists in New York City in the last few years has been sufficient to enable a high degree of selection to be made. During the time in which the experiment was carried on in the commercial concern, in 1917, applications were received from two hundred and eighty-two typists for ninety-six positions which were open during that period.

self by the percentage of best applicants, and the percentage of poorest applicants which it will be able to detect.

Some concerns require an average of twenty-two words of connected discourse to be written per minute, others require thirty-five words, others forty-three words, and so on. Some concerns require twenty words of unconnected discourse to be written per minute, other concerns require ninety-five strokes of indexing work to be written per minute and so on. There are then various degrees of quantitative and qualitative typing required by various concerns and a typist who is capacitated to satisfy the requirements of one concern may not be capacitated to satisfy the requirements of another concern. The tests have developed a definite scale of efficiency in typing, from lower to higher grades. Since each of the various concerns requires a different average ability in its typists, higher or lower than that required by other concerns, it will be necessary, before the tests can be adequately applied, first to determine this required standard for the given concern. Reference can then be made to the scale and a standard time for performance in the tests can be fixed upon. Candidates may then be selected in order, by their grades, from the highest scores down, until the required number have been secured. The employment manager may stop at the "minimum score" in the tests corresponding to the minimum standard of efficiency required for the job and then attempt to procure a new group of applicants in order to secure applicants who can pass the tests with better than minimum grades, or if the supply of possible applicants is limited, he will at least have clear information, at the start, as to what he may expect of those whom he is forced to select from the group who obtain inadequate scores in the tests. Even here he is able to make the best of a bad situation.

APPENDIX

1. The supervisor as a judge of abilities.

In correlating abilities in performing tests with other abilities it is often impossible to obtain objective measures, of individuals in a group, in exact terms of output or of a definite quality and quantity of a particular capacity or ability, in which case measures of abilities are estimated by a teacher or supervisor. The following analysis of a particular supervisor's estimate of abilities may contribute, in some small way, to the reliability which is to be attached to such measures.

The supervisor of the work of groups 2 and 3 was asked to arrange the subjects in each group in an order of merit series for typewriting ability. Emphasis was placed upon the necessity of the supervisor basing his judgments on typewriting ability only and to exclude all other factors. The supervisor was in possession of a copy of all the records of the output. The following table gives the correlations of his judgment with the measures of abilities according to output.

Group 2.		r.
Daily average for five weeks		.74
Daily average for best week		.61
Best day's performance		.59
Group 3.		
Daily average for 5th & 6th weeks		.53
Best day's performance		.42

The correlations tend to show that the supervisor was a better judge of ability which manifested itself over a long rather than a short period of time. He was either a poor judge of ability which manifested itself in spurts rather than ability which manifested itself in steadiness from day to day, or he based his estimates on something else than pure ability, or both alternatives may be true. In conversation with the supervisor, he realized that there were other things which he had taken into consideration in making the judgments, although he tried to judge for actual typing ability only, such as steadiness and general reliability of output, attendance, conduct, ability of the subject to make herself an agreeable member of the group, etc. In other words he was incapable of judging actual typing ability and his estimates were based on the general desirability of the subject to the concern.

This is a problem which every investigator who obtains measures of abilities by means of some one's estimate of those abilities, must confront. It is probably impossible for the general run of supervisors to grade for pure ability and to omit all other factors. This may be a favorable point if tests are sought which will be indices of an individual's total value on a job.

2. Averages of the scores in the tests arranged according to age and education.

In the following tables group M is made up of all the subjects who were tested in the Extension Department of Columbia University and group N is made up of all the subjects who were tested in the industrial concern. The number of subjects in each sub-division is small and the deviations from the average are high, so that, a good amount of caution must be exercised in any interpretation.

When all the evidence is evaluated no difference with age and performance in the tests is found.

Since the comparison, in group M, in education, is made between college graduates, high school graduates, and grammar school graduates, only those subjects who had reached their twenty-second year are used in the group. The tables show that there is a difference between education and performance in some of the tests, but where there is a difference it does not appear to be strongly marked.

58
42

Group 3
Daily average for 5th & 6th weeks
Best day's performance

The correlations tend to show that the supervisor was a better judge of ability which manifested itself over a long rather than a short period of time. He was either a poor judge of ability which manifested itself in spurts rather than in steadiness from day to day, or he based his estimates on something else than pure ability. In conversation with the supervisor, he realized that there were other things which he had taken into consideration in making the judgments, although he tried to judge for actual typing ability only, such as steadiness and general reliability of output, attendance, conduct, ability of the subject to make herself an agreeable member of the group, etc. In other words he was incapable of judging actual typing ability and his estimates were based on the general desirability of the subject to the concern.

SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION 43

AGE

Group M

Age	Subjects	Opposites	Verb-object	Agent-action	Action-agent	Color naming	Hard directions
16	1	35.8	30.6	29.4	44.6	51.0	107.4
17	3	36.33	32.60	42.20	37.07	55.73	109.47
18	14	28.47	31.38	33.54	34.70	55.04	118.10
19	12	31.42	34.28	35.53	35.35	54.03	117.67
20	8	23.68	30.65	29.30	32.45	52.65	110.75
21	8	29.78	28.50	28.65	31.50	56.15	98.45
22-25	10	25.56	28.06	26.74	30.42	50.14	89.86
26-30	9	25.69	33.53	30.18	34.42	56.18	112.04
31-44	12	28.08	36.19	29.92	37.12	52.52	132.98

Group N

16	13	39.78	39.32	43.86	51.52	62.20	176.20
17	40	38.92	36.79	38.35	47.16	62.21	177.40
18	28	43.00	37.62	39.23	51.07	58.13	196.92
19	14	41.13	35.74	35.47	39.79	63.93	195.87
20	5	42.88	37.36	33.72	47.24	65.04	183.72
21	15	44.40	38.88	41.51	48.36	62.33	176.60

Groups M & N combined

16	14	39.50	38.70	42.83	51.03	61.40	171.28
17	43	38.74	36.49	38.62	46.46	61.67	172.66
18	42	38.14	35.54	37.33	45.61	57.10	170.65
19	26	36.65	35.07	35.51	37.77	59.36	159.75
20	13	31.06	33.23	31.00	38.14	57.41	138.82
21	23	39.31	35.27	37.03	42.49	60.19	149.42

EDUCATION

Group N

Education							
Gram. sch. grad.	54	44.51	39.12	41.74	51.03	63.12	186.76
½ yr. high sch.	9	37.40	32.88	35.07	43.87	62.09	190.31
1 yr. high sch.	21	35.81	39.01	35.35	41.43	59.17	187.62
2 yrs. high sch.	19	42.15	36.31	43.25	48.82	62.51	195.14
3 yrs. high sch.	8	35.50	32.05	37.05	50.38	59.82	154.53
4 yrs. high sch.	4	37.70	37.75	37.45	50.20	64.15	158.60
combining,							
Gram. sch. grad.	54	44.51	39.12	41.74	51.03	63.12	186.76
½-1 yr. high sch.	30	36.29	37.17	35.27	42.16	60.05	188.43
2-4 yrs. high sch.	31	39.28	35.39	40.90	49.40	62.01	179.94

Group M

Gram. sch. grad.	8	34.04	30.64	38.84	45.36	59.88	180.08
High sch. grad.	14	26.23	31.30	27.54	33.41	50.93	107.64
College grad.	11	25.11	26.47	25.82	30.87	52.44	95.93

44 *SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION*

3. Sample of an Army Trade Test. (12).*

Welder, Cutter
Cutter, Oxy-acetylene Operator

Oral

- 1. Q. What is it called when the edges of metal sheets are welded together in different spots to hold them in place for welding?
A. Tack (spot). Score 4.
- 2. Q. What chemical is mixed with water to form acetylene gas?
A. Carbide. Score 4.
- 3. Q. What metal do you use to braze brass and cast iron together?
A. Brass (bronze). Score 4.
- 4. Q. What will happen if oil gets on the oxygen regulator or hose?
A. (1) Explode (blow up). Score 4.
(2) Catch fire (burn up). Score 4.

* * * * *

- 18. Q. What is put in the acetylene tank to prevent explosions?
A. Acetone. Score 4.

Rating the candidate

19 and below	N
20 and 21	A+
22 to 35, inclusive	A
36 and 37	A+
38 and 39	J—
40 to 55, inclusive	J
56 and 57	J+
58 and above	E
There is no E— or E+ rating.	

*Vol. 2, p. 148-150.

BIBLIOGRAPHY

1. American Academy of Political Science, Annals of,
Personnel and Employment Problems. 1916. Phila.
Stabilizing Industrial Employment. 1917. Phila.
Steadying Employment. 1916. Phila.
2. Baerwald, Richard. Zur Psychologie der Vorstellungstypen.
Schriften der Gesellschaft für psychologische Forschung.
Leipzig, J. A. Barth, 1916.
3. Boas, Franz. Determination of the Coefficient of Correlation.
Science, N. S., 29, 823-824.
4. Boas, Franz. The Measurement of Variable Quantities. Science
Press, N. Y., 1906.
5. Book, William F. The Psychology of Skill. University of Montana
Studies in Psychology, I, 1908.
6. Brown, H. C. Language and the Association Reflex. Jour. of
Phil. Psy. & Sci. Meth., 13, 1916, 645-649.
7. Brown, William. The Essentials of Mental Measurement. Cam-
bridge University Press, 1911.
8. Bryan, W. L. & Harter, N. Studies in the Physiology and Psy-
chology of the Telegraphic Language. Psy. Rev., 4, 1897,
27-53 & 6, 1899, 346-375.
9. Burt, Harold E. Professor Munsterberg's Vocational Tests. Jour.
of Appl. Psy., I, 1917, 201-213.
10. Chapman, J. Crosby. The Learning Curve in Typewriting. Jour.
of Appl. Psy., 3, 1919, 252-268.
11. Chapman, J. Crosby. Army Trade Tests. Henry Holt & Co., 1921.
12. Chapman, J. Crosby & Hills, Myra E. Positive Acceleration in the
Improvement of a Complex Function. Jour. of Exper.
Psy., I, 1916, 494-507.
13. Chapman, J. Crosby & Toops, Herbert A. A Written Trade Test,
Multiple Choice Method., Jour. of Appl. Psy., 3, 1919,
358-365.
14. Cody, Sherwin. Commercial Tests and How to Use Them. World
Book Co., Yonkers, N. Y., 1919.
15. Committee on the Classification of Personnel in the Army,
Various Members of. How the Army Uses Individual
Differences in Experience. Psy. Bull., 15, 1918, 187-206.
16. Committee on the Classification of Personnel in the Army. The
Personnel System of the United States Army. C. C. P.
339 & 340, Adj. Gen., U. S. A., Wash., D. C., 1919.
17. Frankfurter, Walter. Arbeitsversuche an der Schreibmaschine.
Psychologische Arbeiten, 6, 1914.
18. Garnett, J. C. M. General Ability, Cleverness and Purpose. Bri.
Jour. of Psy., 9, 1919, 345-366.
19. Haggerty, M. T. The Analysis of an Occupation. Manual Train-
ing & Vocational Education, 16, 1915, 472-479.
20. Hart, B. & Spearman, C. General Ability, its Existence and
Nature. Bri. Jour. of Psy., 5, 1912, 51-84.
21. Heinetz, Wilhelm. Vorstudien über die psychologischen Arbeits-
bedingungen des Maschinenschreibens. Zsch. für Angew.
Psy., 13, 1918, 37-90.
22. Herberts, R. Zur Psychologie des Maschinenschreibens. Zsch. für
Angew. Psy., 2, 1909, 551.
23. Heron, David. The Danger of Certain Formulas Suggested for Sub-
stitutes for the Correlation Coefficient. Biometrika, 8, 1911,
109-122.
24. Hollingworth, Harry L. Individual Differences Before, During and
After Practice. Psy. Rev., 21, 1914, 1-8.

46 SOME EMPIRICAL TESTS IN VOCATIONAL SELECTION

25. Hollingworth, Harry L. *Vocational Psychology*. D. Appleton & Co., 1918.
26. Hollingworth, Harry L. & Poffenberger, A. T. *Applied Psychology*, D. Appleton & Co., 1918.
27. Kelly, Truman Lee. *Educational Guidance*. Teachers' College Contributions to Education, 71, 1914.
28. Kelly, Truman Lee. Principles Underlying the Classification of Men. *Jour. of Appl. Psy.*, 3, 1919, 50-67.
29. Kelly, Truman Lee. Tables to Facilitate the Calculation of Partial Coefficients of Correlation and Regression Equations. *University of Texas Bulletin*, 27, 1916.
30. Kemble, William F. Choosing Employees by Mental and Physical Tests. *Engineering Magazine Co.*, N. Y., 1917.
31. Kornhauser, A. W., Meine, F. & Ruml, B. Two Models Showing the Interrelation of Several Significant Correlation Variables. *Psy. Bull.*, 16, 1919, 230-234.
32. Lahy, J. M. Etude Experimentale de l'adaptation Psycho-physiologique aux Actes Volontaires Brefs et Intenses. *Jour. de Psy.*, 10, 1913, 220-236.
33. Link, Henry C. *Employment Psychology*. MacMillan Co., N. Y., 1919.
34. Lipmann, Otto & Stolzenberg, Otto. Methoden zur Auslese Hochwertiger Facharbeitern der Metallindustrie. *Zeit. für Angew. Psy.*, 16, 1920, 173-249.
35. Lough, James E. *Experimental Psychology in Vocational Guidance*. Proceedings of the Second National Conference of the National Vocational Guidance Association, N. Y., 1912.
36. Moore, Henry Ludwell. Forecasting the Yield and Price of Cotton. MacMillan Co., N. Y., 1917.
37. Munsterberg, Hugo. *Psychology and Industrial Efficiency*. Houghton Mifflin, N. Y., 1913.
38. Munsterberg, Hugo. *Vocation and Learning*. People's University, St. Louis, Mo., 1912.
39. National Association of Corporation Schools. *Bulletins* issued Monthly, N. Y., 1911.
40. National Vocational Guidance Association. *Proceedings of the National Conferences*, N. Y.
41. Ogburn, William F. & Goltra, Inez. How Women Vote. *Pol. Sci. Quar.*, 34, 1919, 413-433.
42. Pearson, Karl. On certain errors with regard to multiple correlation occasionally made by those who have not adequately studied the subject. *Biometrika*, Miscellaneous, 10, 1914-1915, 81-87.
43. Pearson, Karl. On the General Theory of the Influence of Selection on Correlation and Variation. *Biometrika*, 8, 1911-1912, 437-443.
44. Pearson, Karl. On some novel properties of partial and multiple correlation coefficients in a universe of manifold characteristics. *Biometrika*, II, 1915-1917, 231-238.
45. Robinson, Edward S. The Analysis of Trade Ability. *Jour. of Appl. Psy.*, 3, 1919, 352-357.
46. Rugg, Harold O. *Statistical Methods Applied to Education*. Houghton Mifflin, N. Y., 1917.
47. Ruml, Beardsley. The Measurement of the Efficiency of Mental Tests. *Psy. Rev.*, 23, 1916, 501-507.
48. Ruml, Beardsley. The Reliability of Mental Tests in the Division of an Academic Group. *Psy. Rev. Mono.*, 24, 1917.
49. Sabatie, J. La Representation Commerciale; notions de psychologie professionnelle et applique. Dunod et Pinart, Paris, 1920.
50. Schneider, Herman. Selecting Young Men for Particular Jobs. *American Machinist*, April, 1913.

51. Soper, H. E. Probable Error of the bi-serial correlation coefficient. *Biometrika*, 10, 1915, 384.
52. Spearman, C. Correlation Calculated from Faulty Data. *Bri. Jour. of Psy.*, 3, 1910, 271-295.
53. Swift, Edgar James. The Acquisition of Skill in Typewriting. *Psy. Bull.*, 1, 1904, 295-305.
54. Swift, Edgar James. Memory of Skillful Movements. *Psy. Bull.*, 3, 1906, 185-187.
55. Swift, Edgar James & Schuyler, William. The Learning Process. *Psy. Bull.*, 4, 1907, 307-310.
56. Thompson, Godfrey H. The Proof or Disproof of General Ability. *Bri. Jour. of Psy.*, 9, 1919, 321-326.
57. Thompson, Godfrey H. The Hierarchy of Abilities. *Bri. Jour. of Psy.*, 9, 1919, 337-344.
58. Thurstone, L. L. The Learning Curve Equation. *Psy. Mono.*, vol. 26, 3, 1919.
59. Thurstone, L. L. Mental Tests for College Entrance. *Jour. of Educ. Psy.*, 10, 1919, 129-142.
60. Thurstone, L. L. A Scoring Method for Mental Tests. *Psy. Bull.*, 10, 1919, 235-239.
61. U. S. Bureau of Education. Pamphlets on Vocational Guidance, Wash., D. C.
62. U. S. Department of Labor. Bulletins issued Monthly, Wash., D. C.
63. Woodworth, Robert S. Combining the Results of Several Tests; a study in statistical method. *Psy. Rev.*, 19, 1912, 97-123.
64. Woodworth, Robert S. & Wells, Frederick L. Association Tests. *Psy. Mono.*, 13, 1911, No. 5, Whole No. 57.
65. Whipple, Guy M. Manual of Mental and Physical Tests. Warwick & York, Baltimore, Md., 1914.
66. Yule, G. Udny. Introduction to the Theory of Statistics. Chas. Griffin & Co., London, 1916.
67. Yule, G. Udny. On the Theory of Correlation. *Jour. of the Roy. Stat. Soc.*, London, 1897, 60, 812-854.

THE JOURNAL OF PHILOSOPHY

There is no similar journal in the field of scientific philosophy. It is identified with no philosophical tradition, and stands preëminently for the correlation of philosophy with the problems and experience of the present.

The contents of recent numbers include

Instinct and Capacity	C. E. AYRES
Behavior and Intelligence	B. H. BODE
The Unity of Consciousness	MAURICE PICARD
Geneticism as a Heuristic Principle in Psychology	JOHN M. FLETCHER
Giving up Instincts in Psychology	ZING YANG KUO
The Intellectual Significance of the Grasping Reflex	R. M. GIVLER
A Tentative Analysis of the Primary Data of Psychology	J. R. KANTOR
Reconstruction in Mental Tests	BEARDSLEY RUMML
The Sanity of Hamlet	TENNEY L. DAVIS
The Meaning of "Religion" and the Place of Mysticism in Religious Life	JAMES H. LEUBA
A Behavioristic View of Purpose	RALPH B. PERRY

Published on alternate Thursdays

SUB-STATION 84, NEW YORK CITY

\$4 PER ANNUM, 26 NUMBERS

20 CENTS PER COPY

ARCHIVES OF PHILOSOPHY

Editorial communications should be addressed to

Professor FREDERICK J. E. WOODBRIDGE, Columbia University,
New York City

Order should be sent to ARCHIVES OF PHILOSOPHY
Sub-Station 84, New York City

The numbers are as follows:

- The Concept of Control. SAVILLA ALICE ELKUS. 40 cents.
 The Will to Believe as a Basis for the Defense of Religious Faith.
 ETTIE STETTMEIER. \$1.
 The Individual: A Metaphysical Inquiry. WILLIAM FORBES COOLEY. \$1.00
 The Ethical Implications of Bergson's Philosophy. UNA BARNARD SAIT. \$1.25.
 Religious Values and Intellectual Consistency. EDWARD H. REISNER. 75 cents.
 Rosmini's Contribution to Ethical Philosophy. JOHN FAVATA BRUNO. 75 cents.
 The Ethics of Euripides. RHYS CARPENTER. 50 cents.
 The Logic of Bergson. GEORGE WILLIAM PECKHAM. 75 cents.
 The Metaphysics of the Supernatural as illustrated by Descartes.
 LINA KAHN. \$1.00.
 Idea and Essence in the Philosophies of Hobbes and Spinoza.
 ALBERT G. A. BALZ. \$1.25.
 The Moral and Political Philosophy of John Locke.
 STERLING TOWER LAMPRECHT. \$1.50.
 Science and Social Progress. HERBERT WALLACE SCHNEIDER. \$1.25.

DIRECTORY OF AMERICAN PSYCHOLOGICAL PERIODICALS

- American Journal of Psychology**—Ithaca, N. Y.; Morrill Hall.
Subscription \$6.50. 600 pages annually. Edited by E. B. Titchener.
Quarterly. General and experimental psychology. Founded 1887.
- Pedagogical Seminary**—Worcester, Mass.: 950 Main Street.
Subscription \$5. 400 pages annually. Edited by G. Stanley Hall.
Quarterly. Pedagogy and educational psychology. Founded 1891.
- Psychological Review**—Princeton, N. J.: Psychological Review Company.
Subscription \$4.25. 480 pages annually.
Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.
- Psychological Bulletin**. Princeton, N. J.: Psychological Review Company.
Subscription \$5. 720 pages annually. Psychological literature.
Monthly. Founded 1904. Edited by Shepherd I. Franz.
- Psychological Monographs**. Princeton, N. J.: Psychological Review Company.
Subscription \$5.50 per vol. 500 pp. Founded 1895. Ed. by James R. Angell.
Published without fixed dates, each issue one or more researches.
- Psychological Index**. Princeton, N. J.: Psychological Review Company.
Subscription \$1.50. 200 pp. Founded 1895. Edited by Madison Bentley.
An annual bibliography of psychological literature.
- Journal of Philosophy**—New York; Sub-Station 84.
Subscription \$4. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge and Wendell T. Bush.
- Archives of Psychology**. Sub-Station 84, New York: Archives of Psychology.
Subscription \$5. 600 pp. ann. Founded 1906. Edited by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.
- Journal of Abnormal Psychology and Social Psychology**—Boston.
Subscription \$5. Richard G. Badger. Edited by Morton Prince.
Bi-Monthly. 480 pages annually. Founded 1906. Abnormal and social.
- Psychological Clinic**. Philadelphia: Psychological Clinic Press.
Subscription \$2.50. 288 pages. Ed. by Lightner Witmer. Founded 1907.
Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.
- Training School Bulletin**. Vineland, N. J.: The Training School.
Subscription \$1. 160 pages annually. Edited by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.
- Journal of Educational Psychology**. Baltimore: Warwick & York.
Subscription \$4. 540 pages annually. Founded 1910.
Monthly (9 numbers). Managing Editor, Harold O. Rugg.
(Educational Psychology Monographs.
Published separately at varying prices. Same publishers.)
- Comparative Psychology Monographs**—Baltimore: Williams & Wilkins Co.
Subscription \$5. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.
- Psychoanalytic Review**. Washington, D. C.: 3617 10th Street, N. W.
Subscription \$6. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Edited by W. A. White and S. E. Jelliffe.
- Journal of Experimental Psychology**—Princeton, N. J.
Psychological Review Company. 480 pages annually. Experimental.
Subscription \$4.25. Founded 1916. Bi-monthly. Ed. by John B. Watson.
- Journal of Applied Psychology**. Worcester, Mass.: Florence Chandler.
Subscription \$4. 400 pages annually. Founded 1917.
Quarterly. Edited by James P. Porter and William F. Book.
- Journal of Comparative Psychology**—Baltimore: Williams & Wilkins Co.
Subscription \$5. 500 pages per volume. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.